

# Nonreactive Research and Secondary Analysis

**Nonreactive Research**  
**Content Analysis**  
**Existing Statistics/Documents**  
**Secondary Analysis of Survey Data**

**Issues of Inference and Theory Testing**  
**Ethical Concerns**  
**Conclusion**

*There are a number of research conditions in which the sole use of the interview or questionnaire leaves unanswerable rival explanations. The purpose of those less popular measurement classes emphasized here is to bolster these weak spots and provide intelligence to evaluate threats to validity. The payout for using these measures is high, but the approach is more demanding of the investigator.*

—Eugene Webb et al., *Nonreactive Measures in the Social Sciences*, pp. 315–316

Behm-Morawitz and Mastro (2008) explored media exposure on young adults, particularly a “mean girls” trend. “Mean girls,” based on a 2004 movie, are teen females who obtain rewards and feel pleasure by being socially aggressive. The authors searched the Internet to identify ninety U.S. teen films released between 1995 and 2005. A teen film stars teen characters and is marketed to a teen audience. From the ninety, they picked the twenty with the highest box office sales. The authors trained three coders for 48 hours on teen films outside the sample. Coders learned to identify primary and secondary characters, socially cooperative behavior (help a friend, resolve conflict), socially aggressive behavior (humiliating others, excluding others), and positive or negative consequences of the behaviors. In data collection, the coders found 139 primary or secondary characters, most (87%) ages 15 to 18, in the twenty films. Slightly more than one-half (55%) were female. Coders identified 337 incidents of socially aggressive behavior and 534 incidents of socially cooperative behavior. In a statistical analysis of the data, the authors found in the films that “both males and females were more often rewarded than punished for engaging in social aggression with females significantly more likely to be rewarded” (p. 136). The authors next conducted a survey of 136 college undergraduates (19–20 years old). They found that the undergraduates who watched the most teen movies and who most identified with teen movie characters were more likely to believe that social aggression is rewarded by increased popularity with peers.

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

Experiments and survey research are both reactive; that is, the people we study are aware of that fact. In this chapter, we look at **nonreactive research**, or research in which the people we study are not aware that they are being studied. We will consider four nonreactive techniques that usually rely on positivist principles but interpretative and critical researchers also use the techniques. We first look at a collection of inventive nonreactive measures and then *content analysis*. *Existing statistics* and *secondary analysis*, the last two techniques, refer to collecting information from government documents or previous surveys. Although the data may have been reactive when first collected, we can address new questions without reactive effects.

### NONREACTIVE RESEARCH

Nonreactive research begins when we notice something that indicates a variable of interest. When we take nonreactive or **unobtrusive measures** (i.e., measures that are not obtrusive or intrusive), the people we study are not aware of it but leave evidence of their social behavior or actions “naturally.” We infer from the evidence to behavior or attitudes without disrupting the people we study. Unnoticed observation is also a type of nonreactive measure, which indicates a construct indirectly.

For example, Rosenbloom et al. (2009) unobtrusively observed and recorded information on 1,062 drivers in two cities (population 300,000), two towns (population 3,000), and two villages (population 800) in Israel. They noted five types of traffic violations: (1) not wearing a seat belt, (2) not using a safety child seat, (3) driving while using a cell phone, (4) failing to comply with a “give way” or yield sign, and (5) stopping in an undesignated area. Based on anonymity in cities they hypothesized that more traffic violations would occur in more urban areas. They found, however, that more traffic violations occurred in towns and villages, and that males committed many more violations than females. The study was nonreactive because the drivers that researchers observed never knew they were part of a study.

### EXAMPLE BOX 1

#### Finding Data on Tombstones

Foster and colleagues (1998) examined the tombstones in ten cemeteries in an area of Illinois for the period 1830 to 1989. They retrieved data on birth and death dates and gender from more than 2,000 of the 2,028 burials. The researchers learned that some trends in the area differed from national ones. They found that conceptions had two peaks (spring and winter), females ages 10 to 64 had a higher death rate than males, and younger people died in late summer but older people in late winter.

### Varieties of Nonreactive or Unobtrusive Observation

Nonreactive measures are varied, and researchers have invented creative ways to measure indirectly social behavior (see Example Box 1, Finding Data on Tombstones). Because the measures have little in common except being nonreactive, we can best learn about them by studying many examples. One type is the **erosion measure**, which considers the wear or deterioration of surfaces. Another is the **accretion measure**, which studies things that have been left behind.<sup>1</sup>

Researchers have examined family portraits in different historical eras to see how gender relations within the family are reflected in seating patterns. Urban anthropologists have examined the contents of garbage dumps to learn about lifestyles from

**Nonreactive research** A type of social research in which people being studied are unaware of the fact.

**Unobtrusive measures** Another name for nonreactive measures that emphasize the fact that the people being studied are not aware of it because the measures do not intrude.

**Erosion measure** Nonreactive measures of the wear or deterioration on surfaces due to the activity of people.

**Accretion measure** Nonreactive measure of the residue of the activity of people or what they leave behind.

**EXPANSION BOX 1****Examples of Nonreactive Measures****PHYSICAL TRACES**

*Erosion:* Wear suggests use.

*Example:* A researcher examines children's toys at a day care that were purchased at the same time. Worn-out toys suggest higher interest in them by the children.

*Accretion:* Accumulation of physical evidence suggests behavior.

*Example:* A researcher examines the brands of aluminum beverage cans in trash or recycling bins in male and female dormitories. This indicates the brands and types of beverages favored by each gender.

**ARCHIVES**

*Running records:* Regularly produced public records may reveal much.

*Example:* A researcher examines marriage records for the bride and groom's ages. Regional differences suggest that the preference for males marrying younger females is higher in certain areas of the country.

*Other records:* Irregular or private records can reveal a lot.

*Example:* A researcher finds the number of reams of paper purchased by a college dean's office for 10 years when student enrollment was stable. A sizable increase suggests that bureaucratic paperwork has increased.

**OBSERVATION**

*External appearance:* How people appear may indicate social factors.

*Example:* A researcher watches students to see whether they are more likely to wear their school's colors and symbols after the school team has won or lost.

*Count behaviors:* Counting how many people do something can be informative.

*Example:* A researcher counts the number of men and women who come to a full stop and those who come to a rolling stop at a stop sign. This suggests gender difference in driving behavior.

*Time duration:* How long people take to do things may indicate their attention.

*Example:* A researcher measures how long men and women pause in front of the painting of a nude man and in front of a painting of a nude woman. Time may indicate embarrassment or interest in same or cross-gender nudity by each gender.

what is thrown away (e.g., liquor bottles indicate level of alcohol consumption). Based on information obtained by their garbage, people underreport their liquor consumption by 40 to 60 percent (Rathje and Murphy, 1992:71). Researchers studied the listening habits of drivers by checking what stations their radios are tuned to when cars are repaired. They measured interest in different museum exhibits by noting worn tiles on the floor in different parts of a museum. They studied differences in graffiti in male versus female high school restrooms to show gender differences in themes. Some researchers examined high school yearbooks to compare the high school activities of those who had psychological problems later in life versus those who did not. Researchers have noted bumper stickers in support of different political candidates

to see whether one candidate's supporters are more likely to obey traffic laws than those of the opposing candidate. Researchers have even measured television-watching habits by noting changes in water pressure due to the use of toilets during television commercials.<sup>2</sup> (Also see Expansion Box 1, Examples of Nonreactive Measures.)

**Recording and Documentation**

Creating a nonreactive measure follows the logic of quantitative measurement, although qualitative researchers also use nonreactive observation. You first conceptualize a construct. Next, you link the construct to a nonreactive measure. The variable's operational definition is how you systematically record observations.

As with other studies, you must rule out reasons for the observation other than the construct of interest. For example, your construct is level of customer interest in ten products displayed in a store. Your operational definition is the amount of customer traffic in front of each of ten store product displays. You measure customer traffic with a hidden video camera. You will need to clarify what the customer traffic means (e.g., the location is near an outside entrance causing more to pass by; people are looking at something beyond the display; people are pausing at the display to put on coats, not looking at the products on it; the floor is a path to another department; or traffic simply indicates a good location for a visual display). Next, you systematically consider what is on the video: Compare it to that in other store locations, look at the number of people at the display, note their speed of walking past or time stopping at the display, and count how many customers turned their heads toward it. You want to record results on a regular basis (e.g., hourly, daily, weekly).

## CONTENT ANALYSIS

### Content Analysis Definition

In a content analysis study, you gather and analyze the content of text. The content can be words, meanings, pictures, symbols, ideas, themes, or any communicated message. The **text** is anything written, visual, or spoken that serves as a medium for communication. It includes books, newspaper or magazine articles, advertisements, speeches, official documents, films or videotapes, musical lyrics, photographs, articles of clothing, Web sites, or works of art. The study about “mean girls” in the box that opened this chapter was an example of a content analysis study. (Also see Example Box 2, What Is the Message of Antiaging Product Web Sites?)

Content analysis has been around for about a century and is used in many fields—literature, history, journalism, political science, education, psychology, and so on. At the first meeting of the

### EXAMPLE BOX 2

#### What Is the Message of Antiaging Product Web Sites?

Ageism, like sexism and racism, requires a set positive or negative stereotypes and messages to reinforce power relations, inequalities, and social privileges. Calasanti (2007) examined Web sites to see what their marketing discourse communicated to older consumers. She identified a sample of 96 antiaging Web sites, coding the pictures and text from each into a set of categories. Coded categories included problems of old age/aging; solutions for problems/old age; gendered aspects of old age; aspects of aging bodies on the site; and depictions of class, race, and sexual orientation. A key message of ageism is that if you can fix your body to forestall aging, you should do so; otherwise, you are a marginal person or loser. Calasanti discovered the antiaging advertisements not only promoted various ways to hide the physical signs of aging but also consistently showed the ideal person as a particular race (White), class (middle class or higher), and sexual orientation (heterosexual). The ads showed men as being dominant in athletic competition or work and sexually assertive. Ads displayed aging women as being alluring sexual partners, competitors with younger women, and sexually receptive to men. The ads suggest that people cannot be old and possess a specific gender at the same time, at least in terms of a White, middle-class, heterosexual ideal. They tell viewers that to look old means losing both gender and sexuality (i.e., becoming a neutral genderless person who is neither male nor female) and that only by appearing younger can they restore these socially valued assets.

German Sociological Society, in 1910, Max Weber suggested using it to study newspapers.<sup>3</sup> In quantitative content analysis, you use objective and systematic counting and recording procedures to produce a numerical description of the content in a text. There are also qualitative or interpretive versions of

**Text** A general name for a communication medium from which symbolic meaning is measured in content analysis.

**EXPANSION BOX 2****How Qualitative Researchers Study Documents or Statistical Reports**

Qualitative researchers who use interpretative or critical approaches also study documents and reports with statistical information, but they tend to do so differently from positivists. They consider documents and statistical reports to be cultural objects, or media that communicate social meaning. They see the documents as belonging to a range of other cultural objects (e.g., monuments, diaries, musical scores, shopping lists, films, photographs, paintings, engineering drawings, Web pages) that carry meaning. For example, an architectural floor plan is a document that expresses spatial arrangements that convey social meanings. Some offices are located in desirable locations with large windows designed for holders of certain highly ranked job positions.

Instead of treating a document or statistical report as a neutral container of content, qualitative researchers examine the larger context of its creation, distribution, and reception. Consistent with a constructionist perspective, qualitative researchers emphasize the entire process from a document's creation (including the intentions of creators) through its consumption or reception by various receivers/consumers and then situate the document in a social context. In short, they treat the document or report as a cultural object that carries social meaning in its own right. Although they may examine the content of a document or report, they do not limit themselves to it.

Qualitative researchers emphasize that people think and interact on the basis of meaning as well as with words or numbers. For example, the content in one document may convey medical information to health care workers, grant a person access to a social service, sell products to a consumer, inform officials of geographic areas where problems exist, or allow/prevent a person's entry into a country. Different people may put the same document or report to different uses at different times, and processes of "reading" or interpreting documents often depend on training and following rules. For example, people learn what to look for in a medical record, statistical report, or passport. People looking at the same document may see different things, follow different rules, and use it for different purposes (e.g., grant insurance reimbursement or prescribe a medical treatment, test a hypothesis or allocate funds for a new public building, allow someone into a country, or cash a check). Qualitative researchers look at multiple facets of a document and its content. For example, a magazine article can carry content that entertains readers, is a vehicle that allows an author to build a reputation, triggers a public controversy, and is a way to boost magazine sales; (see Griswold (1987, 1994) and Prior (2003) on the study of cultural objects and documents).

content analysis (see Expansion Box 2, How Qualitative Researchers Study Documents or Statistical Reports). Here the focus is on quantitative data about a text's content.

Content analysis is nonreactive because the process of placing words, messages, or symbols in a text to communicate to a reader or receiver is without any awareness of the researcher. For example, I, as author of this book, wrote words or drew diagrams to communicate research method content to you, the student. The way I wrote this book and the way you read it are without any knowledge or intention of its ever being content analyzed.

Content analysis lets you see and reveal the content (i.e., messages, meanings, symbols) within a communication source (i.e., a book, article,

movie). You probe into and discover content in a manner different from the ordinary way of reading a book or watching a television program. Content analysis can document—in objective, quantitative terms—whether feelings based on unsystematic observation are true. It yields repeatable, precise results about the text. After you gather the data, you analyze them with statistics in the same way that an experimenter or survey researcher would.

**Topics Appropriate for Content Analysis**

Content analysis is used for many purposes: to study themes in popular songs and religious symbols in hymns, trends in the topics that newspapers cover and the ideological tone of newspaper editorials,

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

gender role stereotypes in textbooks or feature films, frequency with which people of different races appear in television commercials and programs, answers to open-ended survey questions, enemy propaganda during wartime, the covers of popular magazines, personality characteristics from suicide notes, themes in advertising messages, gender differences in conversations, and so on.

Generalizations you make on the basis of content analysis are limited to the cultural communication itself. Content analysis cannot determine the truthfulness of an assertion or evaluate the aesthetic qualities of literature. It reveals the content in text but cannot interpret the content's significance. You should examine the text directly. Holsti (1968a:602) warned, "Content analysis may be considered as a supplement to, not as a substitute for, subjective examination of documents."

Content analysis is useful for three types of research questions: those regarding a large volume of text, content that may be at a distance or scattered, and content that is difficult to see or document with casual observation. You can measure large amounts of text (e.g., 20 years of newspaper articles) with sampling and multiple coders. You can study topics "at a distance" such as broadcasts in a hostile foreign country or scattered such as common themes in fifteen films produced by the same director over a 20-year period. Most important, content analysis can reveal messages in a text that are difficult to see with casual observation. Even the creator of the text or those who read it may be unaware of all its themes, biases, or characteristics. For example, authors of preschool picture books may not consciously intend to portray children in traditional stereotyped gender roles, but a high degree of such stereotyping has been revealed through content analysis.<sup>4</sup> Another example is that of conversations in all-male versus all-female groups. Although people may be unaware of it, in same-gender groups, women talk more about interpersonal matters and social relationships whereas men talk more about achievement and aggressive themes.<sup>5</sup>

### Measurement and Coding

As in most quantitative research, careful measurement is crucial in content analysis. You take diffuse

and murky symbolic communication and convert it into precise, objective, quantitative data. To do this, you must very carefully design and document procedures for coding to make replication possible. For example, you want to determine how frequently television dramas portray elderly characters in terms of negative stereotypes. You must develop a measure of the construct "negative stereotypes of the elderly." The conceptualization may be a list of stereotypes or negative generalizations about older people (e.g., senile, forgetful, cranky, frail, hard of hearing, slow, ill, in nursing homes, inactive, conservative) that may or may not accurately reflect elderly people. For example, if 5 percent of people over age 65 are in nursing homes yet 50 percent of those over age 65 on television shows are portrayed as being in nursing homes, evidence supports negative stereotyping.<sup>6</sup>

In a content analysis study, you operationalize constructs with a **coding system**. It is a set of instructions or rules describing how to observe and record content from text. You tailor it to the type of text or communication medium you are studying (e.g., television drama, novels, photos in magazine advertisements). It also depends on your unit of analysis.

The unit of analysis can vary a great deal in content analysis. It can be a word, a phrase, a theme, a plot, a newspaper article, a character, and so forth. In the study on "mean girls" in this chapter's opening box, the unit of analysis was film characters. In addition to units of analysis, you use other units in content analysis that may or may not be the same as units of analysis: recording units, context units, and enumeration units. There are few differences among them, and they are easily confused, but each has a distinct role. In simple projects, all three are the same. For example, you may note features of television commercials for cars or trucks (commercial is recording unit) and what television show or other commercial appeared before or after it (context unit) and count the number and features of people

**Coding system** A set of instructions or rules used in content analysis to explain how a researcher systematically converted the symbolic content from text into quantitative data.

appearing in each commercial (person is an enumeration unit).

Measurement in content analysis uses **structured observation**: systematic, careful observation based on written rules. The rules explain how to categorize and classify observations. As with other measurement, categories should be mutually exclusive and exhaustive. Written rules make replication possible and improve reliability. Although researchers begin with preliminary coding rules, they often conduct a pilot study and refine coding based on it. Coding systems identify four characteristics of text content: frequency, direction, intensity, and space. A researcher measures from one to all four characteristics in a content analysis research project (see Expansion Box 3, What We Measure).

### Coding, Validity, and Reliability

Coding requires carefully looking at text and converting it in a very systematic manner into measures of significant words, symbols, or messages. There are two major types of content analysis coding: manifest and latent.

**Manifest coding** involves the visible, surface content in a text. For example, you count the number of times a phrase or word (e.g., red) appears in written text or whether a specific action (e.g., a kiss) appears in a video scene. The manifest coding system has a list of terms or actions that you want to locate. For written words, you can scan the information into an electronic form and use a computer program to search for words or phrases and let a computer count the number of times they appear. To do this, you first create a comprehensive list of relevant words or phrases.<sup>7</sup>

**Structured observation** A method of watching what is happening in a social setting that is highly organized and follows systematic rules for observation and documentation.

**Manifest coding** A type of content analysis coding in which a researcher first develops a list of words, phrases, or symbols and then locates them in a communication medium.

### EXPANSION BOX 3

#### What We Measure

1. *Frequency* refers to whether or not something occurs and, if it occurs, how often. For example, how many elderly people appear on a television program within a given week? What percentage of all characters are they, or in what percentage of programs do they appear?
2. *Direction* refers to the direction of messages in the content along some continuum (e.g., positive or negative, supporting or opposing). For example, we devise a list of ways an elderly television character can act. We classify the actions into three categories: positive (e.g., friendly, wise, kind, considerate), neutral, or negative (e.g., nasty, dull, selfish, slow, forgetful).
3. *Intensity* is the strength or power of a message in a direction. A television character may be active, (e.g., running about, speaking quickly and loudly) or passive (e.g., standing nearly still and saying a few words quietly). A characteristic, such as forgetfulness, can be minor (e.g., not remembering to take car keys when leaving home, taking longer time to recall the name of someone who has not been seen in 10 years) or major (e.g., not remembering one's own name, not recognizing one's children).
4. *Space* is the size of a text message, amount of time, or the amount of space allocated to a message. It is easy to measure size or space of a print advertisement or a photo. We can measure space in written text by counting words, sentences, paragraphs, or the space it covers on a page (e.g., square inches). For video or audio text, we measure the amount of time allocated. For example, a TV character may be present for a few seconds or in every scene of an hour-long program.

Manifest coding is highly reliable because the phrase or word either is or is not present. Unfortunately, manifest coding does not consider the connotations of words or phrases. The same word can take on different meanings depending on the context. The possibility that there are multiple meanings of a word limits the measurement validity of manifest coding.

For example, I read a book with a *red* cover that is a real *red* herring. Unfortunately, its publisher

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

drowned in *red* ink because the editor could not deal with the *red* tape that occurs when a book is *red* hot. The book has a story about a *red* fire truck that stops at *red* lights only after the leaves turn *red*. There is also a group of *Reds* who carry *red* flags to the little *red* schoolhouse. They are opposed by *red*-blooded *rednecks* who eat *red* meat and honor the *red*, white, and blue. The main character is a *red*-nosed matorador who fights *red* foxes, not bulls, with his *red* cape. *Red*-lipped Little *Red* Riding Hood is also in the book. She develops *red* eyes and becomes *red*-faced after eating a lot of *red* peppers in the *red* light district. She is given a *red* backside by her angry mother, a *red*head.

**Latent coding** (also called *semantic analysis*) looks for the underlying, implicit meaning in the content of a text. For example, you read an entire paragraph and decide whether it contains erotic themes or a romantic mood. Your coding system contains general rules to guide your interpretation of the text and to determine whether particular themes or moods are present. The study on “mean girls” in the chapter’s opening box used latent coding, which tends to be less reliable than manifest coding. It depends on a coder’s knowledge of language and social meaning.<sup>8</sup> Training, practice, and written rules improve reliability, but still it is difficult to consistently identify themes, moods, and the like. However, the validity of latent coding can exceed that of manifest coding because we communicate meaning in many implicit ways that depend on context, not just specific words.

You may want to use both manifest and latent coding to study the content of text. Agreement from the two approaches strengthens your final result; if they disagree, you should reexamine the operational and theoretical definitions.

In many studies, you will need to code information from a very large number of units. You might look at the content in thirty books, hundreds of hours of television programming, or about one hundred Web sites (as in the opening box). In addition to coding the information personally, you may hire assistants to help with the coding. You teach coders the coding system and train them to fill out a recording sheet. Coders should understand the variables, follow the coding system, and ask about

ambiguities. You must record all decisions about how to treat a new specific coding situation after coding begins so that you can be consistent.

If you use several coders, you must always check for consistency across coders. To do this, you ask coders to code the same text independently and then check for consistency across coders. You measure **intercoder reliability**, a type of equivalence reliability, with a statistical coefficient that identifies the degree of consistency among coders (see Expansion Box 4, Krippendorff’s Alpha).<sup>9</sup> You always report the coefficient with the results of content analysis research. The study described in the chapter’s opening box reported an intercoder reliability measure (*Krippendorff’s alpha*) for each variable measured. To create the coefficient, the three coders each coded 10 percent of all the films used in the study. The alpha coefficient ranged from 0.72 to 1.0, with most over 0.80.

If the coding process stretches over considerable time (e.g., more than 3 months), you should also check stability reliability by having each coder independently code samples of text that were previously coded to see whether the coding is stable or changing. For example, you have 6 hours of television episodes coded in April. You ask the coders to code the 6 hours again in September without allowing the coders to look at their original coding decisions. If the results are the same, you have stability reliability. If you see large deviations in coding, you may need to retrain coders and recode a second time.

Researchers have studied many forms of visual “text,” such as photographs, paintings, statues, buildings, clothing, videos, and film. Visual “text” is more difficult to analyze than written text because it communicates messages or emotional content indirectly through images, symbols, and metaphors.

**Latent coding** A type of content analysis coding in which a researcher identifies subjective meaning such as themes or motifs and then systematically locates them in a communication medium.

**Intercoder reliability** Equivalence reliability in content analysis with multiple content coders that requires a high degree of consistency across coders.



**EXPANSION BOX 4****Krippendorff's Alpha**

Krippendorff's alpha ( $\alpha$ ) is the most widely used and best known measure of intercoder agreement or interrater reliability. Klaus Krippendorff developed this intercoder reliability coefficient to measure the agreement between observers, coders, judges, raters, and measuring instruments. When observers agree perfectly, observed disagreement of  $\alpha = 1$  and indicates perfect reliability. Agreement by observers as if chance had produced the results indicates the absence of reliability,  $\alpha = 0$ . It is as if the coders failed to observe the text or information and made up their data by throwing dice.

$$\alpha\text{'s general form is: } \alpha = 1 - \frac{D_o}{D_e}$$

where  $D_o$  is the observed disagreement and  $D_e$  is the disagreement one would expect when the coding of units is attributable to chance rather than to the properties of these units.

The mathematics behind the formula and its more advanced details are beyond the level of this book (see Hayes and Krippendorff, 2007, and Krippendorff, 2004). The data for this formula come from two or more jointly trained coders working independently to assign values to a variable for a common set of units of analysis. Details of the coefficient will change based on the number of coders, range of values in variables, and so forth. The coefficient  $\alpha$  applies to many situations: any number of coders, any number of variable categories or measures, any level of measurement (nominal, ordinal, interval, ratio), any incomplete or missing data, and any sample size. Several statistical computer programs can compute the statistic.

Moreover, visual images often contain mixed messages and operate at multiple levels of meaning.

Most people share a common meaning for key symbols of the dominant culture, but people may read a symbol differently. For example, should you "read" a nation's flag to mean patriotism, duty to nation, and honor of tradition, or domination, abuse of power, and police or military aggression? Japan rarely displayed its national flag in public schools from 1945 to 1999. Then government officials

enacted a law that required its display and the playing of the national anthem, causing great controversy. Conservative politicians wanted the flag displayed to instill more patriotism among the nation's youth. However, many teachers and others objected because of the flag was strongly associated with Japan's past military aggression and suppression of democracy, and extremist right-wing groups in Japan often promoted the display of the flag.

The confederate flag in the United States contains sharply divergent meanings for different social groups.<sup>10</sup> To many African Americans, it symbolizes racial segregation, slavery, and violent oppression by Whites during the Jim Crow era. For many older Whites, it symbolizes regional heritage and a genteel "Old-South" lifestyle. For others, it symbolizes rebelliousness, individual freedom, and rejection of externally imposed authority. For some people outside the United States, it is simply a colorful fashion statement with connections to the United States. There are several possible readings of the flag as a symbol.

To study visual images, you must learn to "read" multiple meanings of visual text and to interpret various symbolic images. Such a "reading" is not mechanical (i.e., image X always means G) but depends on the cultural context because the meaning of an image is culture bound. It also depends on the interrelationships within a field of many symbols. The meaning of the confederate flag may vary by age, racial group, geographic location, and so forth. It also varies by how it is displayed. Displaying the flag at a Klu Klux Klan rally, at a University of Mississippi football game, as part of the Georgia state flag, and on the back of a motorcycle "biker" jacket may not carry the same meaning. In my hometown, I read a newspaper article stating that the police are tracking a high school gang advocating racial hate that fights and intimidates non-Whites. The gang symbol (on hats and jackets) is the confederate flag. National symbols, such as the Statue of Liberty, are also used to convey social or political messages (see Example Box 3, Magazine Covers and Immigration).

Sociopolitical groups construct new symbols or wrestle for control of the meaning of major existing symbols. For example, some people want

**EXAMPLE BOX 3****Magazine Covers and Immigration**

Chavez (2001) conducted a content analysis of the covers of major U.S. magazines that dealt with the issue of immigration into the country. Looking at the covers of ten magazines from the mid-1970s to the mid-1990s, he classified the covers as having one of three major messages: affirmative, alarmist, or neutral or balanced. Beyond his classification and identifying trends in messages, he noted how the mix of people (i.e., race, gender, age, and dress) in the photographs and the recurrent use of major symbols, such as the Statue of Liberty or the U.S. flag, communicated messages. Chavez argued that magazine covers are a site, or location, where cultural meaning is created. Visual images on magazine covers have multiple levels of meaning, and viewers construct specific meanings as they read the image and use their cultural knowledge. Collectively, the covers convey a worldview and express messages about a nation and its people. For example, a magazine cover that displayed the icon of the Statue of Liberty as strong and full of compassion (message: welcome immigrants) was altered to have strong Asian facial features (message: Asian immigrants distorted the national culture and altered the nation's racial makeup), or holding a large stop sign (message: go away immigrants). Chavez (p. 44) observed that "images on magazines both refer to and, in the process, help to structure and construct contemporary 'American' identity."

to assign a religious meaning to the Christmas tree; others want it to represent a celebration of tradition and family values without religious content; still others want it to mean a festive holiday season for commercial reasons. Because of the complex, multilayered meanings of symbols, you need to combine qualitative judgments about the images with quantitative data in content analysis.

**How to Conduct a Content Analysis Study**

1. *Formulate the research question.* You begin with a topic and a research question. When the question involves variables that are messages or symbols, content analysis may be appropriate. For

example, you want to study how local television covers a campaign for mayor of the city. Your question may be whether each candidate has equal coverage. The construct "coverage" includes the amount of coverage (time on television), the prominence of the coverage, and whether the coverage favors one candidate over another. You could survey people and ask what they think of the coverage, but a better strategy is to examine the news reports directly using content analysis.

2. *Decide on units of analysis.* You must decide on the units of analysis. For example, for a political campaign, each day of a news show on each of several local stations could be your unit of analysis, or each news report or segment during each of two evening news programs each day on all local stations. You could also count television advertisements by candidates or issue groups. You could study debate or interview programs on television that featured the candidates.

3. *Develop a sampling plan.* Random sampling is very useful in content analysis. First, you must define the population and the sampling element. For example, the population might be all words, all sentences, all paragraphs, or all articles in certain types of documents over a specified period. Likewise, it could include each conversation, situation, scene, or episode of certain types of television programs over a particular period. For example, you may want to see how the candidates are covered on television news programs, commercials, and in debate or interview programs during the one year leading up to the election and the month following it. You must decide whether to include news programs during the daytime and special reports on just two Monday through Friday evening news programs. Should you include commercials aired any time of the day any day of the week or limit your population to times when more people view television? Your unit of analysis could be the news program segment that focuses on the campaign and names a candidate, a commercial in which a candidate's image or name appears, or the interview program featuring one or more candidates. Your population may include all news program segments, commercials, and interview programs aired on four local television stations during a 13-month period.

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

After you have specified the target population and your sampling elements, you will need to plan in detail. This includes constructing sampling frames and determining sample size and the sampling ratios. The sampling frame is a critical step in creating an accurate random sample. To sample coverage of mayoral candidates over a 13-month period, you should list all news segments, commercials, and debate-interview programs. Practically, you might have three separate samples, one for commercials, one for news segments, and one for debates. Because the same commercial could appear many times, you might create a list (sampling frame) of all possible time slots when commercials could appear and randomly sample the time slots. Because a news program could have none, one, or several segments focusing the election and candidates, you may have to search each station's program log to obtain a list of all possible segments and then use that as a sampling frame. If there are only a few debate-interview programs, you might include the entire population.

As you plan a project, you should calculate the work required. For example, during a pilot test, you might find that it takes an average of 15 minutes to view and code a 30- or 60-second commercial, 20 minutes for a 3- to 5-minute news program segment, and 2 hours for a 30-minute debate or interview program. This does not include time for sampling or locating the commercial, segment, or debate. Let us say the sampling frame had 300 commercials, 80 news segments, and five debates, and you sampled 100 of the commercials, 40 segments, and all five debates. Your coding time would be  $(15 \text{ minutes} \times 100 \text{ commercials}) + (20 \text{ minutes} \times 40 \text{ segments}) + (180 \text{ minutes} \times 5 \text{ debates}) = 2,200 \text{ minutes}$ , or about 37 hours after you have gathered and organized all of the video feeds. You might consider hiring assistants as coders.

4. *Construct coding categories and a recording sheet.* You need to identify all variables of interest. Often they will come from ideas in a literature review, from your own thinking or theory, or from a preliminary analysis of pilot data. You should create a very explicit coding system for yourself and for coders if you use them. The manifest or latent coding system will describe exactly how to convert what a coder sees

or hears into a few code categories. To organize codes, you should create a recording sheet. This is a grid or page with a place to record the identification number of the unit and spaces for coding information about each variable (see Figure 1). You should always test your coding system and recording sheets with some pilot data (about a dozen units).

5. *Coding and intercoder reliability check.* Finally, if you use multiple coders, check intercoder reliability. Usually this means selecting 10 percent of your total sample and having each coder use the coding system with the same units but independently of one another. If necessary, discard and recode information for inaccurate coders.

6. *Data collection and analysis.* After you have prepared the coding system and recording sheets and trained all coders, you are ready to gather and check the data. You enter the data into a computer for statistical analysis, interpret the results, and prepare a report.

### Inferences

The inferences that you can make based on the results is a critical issue in content analysis, which describes what is in the text. It cannot reveal the intentions of those who created the text or the effects that messages in the text have on those who receive them. For example, content analysis shows that children's books contain gender stereotypes. That does not necessarily mean that the stereotypes in the books shape the beliefs or behaviors of children; you need to conduct a separate study on children's perceptions to verify that inference. In the study described in this chapter's opening box, the authors conducted a second study with the survey method to see how their content analysis results affected the viewers of teen movies.

## EXISTING STATISTICS/DOCUMENTS

### Appropriate Topics

Many types of information about the social world are already available in the form of statistical documents (books, reports, etc.) or as published compilations available in libraries or on computerized records. In either case, you can search through

NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

**FIGURE 1 Example TV Commercial Recording Sheet for Content Analysis on Mayoral Election**

Commercial ID \_\_\_\_ Date of commercial \_\_\_\_ Time of commercial \_\_\_\_ Coder # \_\_\_\_  
Station on which commercial aired \_\_\_\_ Duration \_\_\_\_ Sponsor of commercial \_\_\_\_  
Visual of Candidate A shown? \_\_\_\_ Voice of Candidate A heard? \_\_\_\_  
Visual of Candidate B shown? \_\_\_\_ Voice of Candidate B heard? \_\_\_\_  
Visual of Candidate C shown? \_\_\_\_ Voice of Candidate C heard? \_\_\_\_  
Number of people shown other than the candidates \_\_\_\_  
Number who appear elderly \_\_\_\_ Number of nonelderly adults \_\_\_\_ Number of children \_\_\_\_  
Gender mix \_\_\_\_ Racial-ethnic mix W\_\_ B\_\_ H\_\_ A\_\_ O\_\_  
Occupational mix \_\_\_\_\_

Check all issues referenced in commercial that apply.

1. Taxes \_\_\_\_
2. Schools \_\_\_\_
3. Crime \_\_\_\_
4. Housing \_\_\_\_
5. Public transport \_\_\_\_
6. Roads \_\_\_\_
7. Social services \_\_\_\_
8. Parks \_\_\_\_
9. Business \_\_\_\_
10. Urban sprawl \_\_\_\_
11. Water/air quality \_\_\_\_
12. Vision for future \_\_\_\_
13. Neighborhoods \_\_\_\_
14. Fire/police \_\_\_\_
15. Zoning \_\_\_\_
16. Library \_\_\_\_
17. Quality of life \_\_\_\_
18. Elderly services \_\_\_\_
19. Youth services \_\_\_\_
20. Multicultural issues \_\_\_\_
21. Public health \_\_\_\_
22. Job creation \_\_\_\_
23. Efficient government \_\_\_\_
24. Public voice \_\_\_\_
25. Cooperation with other government agencies \_\_\_\_

Endorsements of candidate A made by \_\_\_\_\_

Endorsements of candidate A made by \_\_\_\_\_

Endorsements of candidate A made by \_\_\_\_\_

Criticism of candidate A made by \_\_\_\_\_ on issue number above \_\_\_\_

Criticism of candidate B made by \_\_\_\_\_ on issue number above \_\_\_\_

Criticism of candidate C made by \_\_\_\_\_ on issue number above \_\_\_\_

such collections of information with a research question and variables in mind and then reassemble and statistically analyze the information in new ways to address a research question.

It is difficult to specify topics that are appropriate for existing statistics research because they are so varied. You can study any topic on which an organization collected information and made it publicly available. In fact, existing statistics projects do not fit neatly into a deductive model of research design. Rather, you creatively reorganize the existing information into the variables for a research question after first finding what data are available.

Recall that experiments are best for topics that can be controlled and manipulated as independent variables. Survey research is best for topics about which we ask questions to learn about reported attitudes or behaviors. Content analysis is best for topics that involve the content of messages in cultural communication. Existing statistics research is best for topics that involve information collected by large bureaucratic organizations. Public or private organizations systematically gather many types of information for policy decisions or as a public service. Rarely do they collect data for purposes directly related to a specific research question. Thus, existing statistics research is appropriate for testing hypotheses that involve variables in official reports of social, economic, and political conditions. These include descriptions of organizations or the people in them. Often organizations collect the information over long time periods. For example, you can use existing statistics to see whether unemployment and crime rates are associated in 150 cities across a 20-year period.

Existing statistics are valuable for looking over time and across nations. Recall the existing statistics study about red and blue states by McVeigh and Sobolewski (2007). The census (see Expansion Box 5, The Census) is a valuable type of existing statistical data (see Example Box 4, Existing Census Statistics and Naturalization in the Early Twentieth Century).

### Social Indicators

During the 1960s, many social scientists, dissatisfied with the information available to decision

## EXPANSION BOX 5

### The Census

Almost every country conducts a census, or a regular count of its population. For example, Australia has done so since 1881, Canada since 1871, and the United States since 1790. Most nations conduct a census every 5 or 10 years. In addition to the number of people, census officials collect information on topics such as housing conditions, ethnicity, religious affiliation, education, and so forth.

The census is a major source of high-quality existing statistical data, but it can be controversial. In Canada, an attempt to count the number of same-sex couples living together evoked public debate about whether the government should document the changes in society. In Great Britain, the Muslim minority welcomed questions about religion in the 2001 census because they felt that they had been officially ignored. In the United States, the measurement of race and ethnicity was hotly debated, so in the 2000 census, people could place themselves in multiple racial-ethnic categories.

The U.S. 2000 census also generated a serious public controversy because it missed thousands of people, most from low-income areas with concentrations of recent immigrants and racial minorities. Some double counting of people in high-income areas where many owned second homes also occurred. A contentious debate arose among politicians to end miscounts by using scientific sampling and adjusting the census. The politicians proved to be less concerned about improving the scientific accuracy of the census than retaining traditional census methods that would benefit their own political fortunes or help their constituencies because the government uses census data to draw voting districts and allocate public funds to areas.

makers, spawned the “social indicators’ movement” to measure social well-being. They wanted to expand understanding by combining information about social well-being with generally used indicators of economic performance (e.g., gross national product) to better inform policy-making officials. Members of this movement hoped that measuring the quality of social life would influence public policy decisions.<sup>11</sup> Today, many books, articles,

**EXAMPLE BOX 4****Existing Census Statistics and Naturalization in the Early Twentieth Century**

Bloemraad (2006) studied existing statistical records to examine citizenship acquisition, or naturalization, in early twentieth century America. She noted that many commentators contrast low levels of citizenship acquisition among today's immigrants with the assumed rapid and uniform naturalization of European migrants 80 to 100 years ago. However, there is little solid evidence about the earlier process, and myths have filled the void. Bloemraad examined data on adult male immigrants from the 1900, 1910, and 1920 U.S. censuses. The U.S. government gathers census data and makes them available to the public for statistical analysis. Naturalization is a legal process that enables noncitizens to become citizens. It requires a specific length of residence, a clean legal record, passage of a language test, and several other features. Between 1900 and 1920, the proportion of immigrants who held U.S. citizenship fell from 67 percent to 49 percent due to large-scale immigration from Europe. Bloemraad used sophisticated statistical analysis to

investigate four explanations for naturalization: individuals' resources and skills, regulatory and bureaucratic barriers to citizenship, relative costs and benefits of citizenship, and the degree of political mobilization directed to immigrants. The census had measures of years of residence, age, literacy, and English ability. Bloemraad found that naturalization rates varied widely by geographic area. The key factor that made a difference was the warmth of the welcome extended to newcomers. Although some local histories suggested this process, we did not have the nationwide, generalizable findings until her large-scale national study that statistically analyzed millions of cases. Her most notable result was that where an immigrant lived had a more significant effect on naturalization than the immigrant's birthplace, ability to speak English, or literacy. In short, the local reception of immigrants determined how fast they became citizens, not features of the individual immigrants as was emphasized in past studies or by political commentators.

and reports discuss social indicators. A scholarly journal, *Social Indicators Research*, is devoted to the creation and evaluation of social indicators. Since 1976 every three years, the U.S. Census Bureau has published a report, *Social Indicators*, and the United Nations collects many measures of social well-being across nations.

A **social indicator** is any measure of social well-being that can inform policy decisions. Many specific indicators can measure well-being related to the following areas: population, family, housing, social security and welfare, health and nutrition, public safety, education and training, work, income, culture and leisure, social mobility, voting, and participation in social and religious organizations. The FBI's uniform crime index indicates the amount of crime in U.S. society. Social indicators can measure negative aspects of social life, such as the infant mortality rate (the death rate of infants during the first year of life) and alcoholism, or they can indicate positive aspects, such as job satisfaction or the percentage of housing units with indoor

plumbing. Social indicators often involve implicit value judgments (e.g., which crimes are serious or what constitutes a good quality of life).

The Institute for Innovation in Social Policy now at Vassar College created an Index of Social Well-Being for the United States. It combines measures of sixteen social problem areas (see Chart 1) from various existing U.S. government statistical documents. With it you can compare each year to the best level recorded for an item, on a scale of 0 to 100 with 100 being the highest score, since 1970 when the index began. The United States reached its highest level of social well-being in 1973 (index score = 77.5) and has since declined. Overall, between 1970 and 2007, the Index declined from 66 to 56. The current social well-being is lower than in the recent past and varies greatly by state.<sup>12</sup>

**Social indicator** A quantitative indicator of social well-being.

NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

**CHART 1 Social Health Index of United States**

**SIXTEEN SOCIAL INDICATORS USED TO CREATE SOCIAL HEALTH INDEX**

Infant mortality	Health insurance coverage
Child poverty	Aging: poverty among the elderly
Child abuse	Suicide among the elderly
Teenage suicide	Homicide
Teenage drug abuse	Alcohol-related traffic fatalities
High school completion	Food stamp coverage
Unemployment	Affordable housing
Average wages	Income inequality

**Social Health of States in 2008**

<b>RANK</b>	<b>SOCIAL HEALTH SCORE</b>	<b>RANK</b>	<b>SOCIAL HEALTH SCORE</b>
1. Minnesota	75.0	26. Missouri	51.4
2. Iowa	71.1	27. Michigan	48.9
3. New Hampshire	67.2	28. Oregon	47.8
4. Nebraska	67.0	29. Rhode Island	46.8
5. Hawaii	63.1	30. Colorado	44.6
6. Vermont	62.7	31. New York	43.9
7. Connecticut	61.2	32. Georgia	43.7
8. North Dakota	61.1	33. Alaska	43.6
9. Utah	60.6	34. Nevada	42.6
10. New Jersey	59.9	35. California	41.7
11. Idaho	59.7	36. West Virginia	40.8
12. Virginia	59.7	37. Oklahoma	40.1
13. Pennsylvania	58.6	38. Montana	39.4
14. Maine	57.4	39. Alabama	38.8
15. Indiana	55.9	40. South Carolina	38.0
16. Kansas	55.9	41. Texas	37.8
17. Delaware	55.7	42. Louisiana	37.5
18. Illinois	55.2	43. Arkansas	36.4
19. Wisconsin	55.2	44. Kentucky	36.2
20. Maryland	54.9	45. Tennessee	35.5
21. South Dakota	54.4	46. Florida	34.3
22. Ohio	53.8	47. North Carolina	33.4
23. Wyoming	53.4	48. Arizona	32.8
24. Massachusetts	53.1	49. Mississippi	31.0
25. Washington	52.2	50. New Mexico	26.8

Source: From <http://iisp.vassar.edu/socialhealth08.html>. Institute for Innovation in Social Policy. Vassar College. Reprinted by permission.

### Locating Data

Government or international agencies and private sources are the main providers of existing statistics. If you plan to conduct an existing statistics study, it is wise to discuss your interests with an information professional—in this case, a reference librarian, who can direct you to possible sources. Most existing documents are “free”—that is, available at public libraries—but the time and effort required to search for specific information can be substantial. Researchers who conduct existing statistics research spend many hours in libraries or on the Internet. After they locate the information, they record it on computer files or recording sheets for later analysis. Often it is already available in an electronic format. For example, instead of recording voting data from reference books, researchers might use a social science data archive at the University of Michigan (to be discussed). Also see Expansion Box 6, Newspaper Reports as a Data Source.

Researchers can be very creative using existing statistics. (See Example Box 5, Existing Statistics, Androgynous First Names, and Collective Behavior.) With many sources available, I will discuss only a small sample of them here. The

### EXPANSION BOX 6

#### Newspaper Reports as a Data Source

Many social researchers use reports in newspapers as a data source, not only to analyze the content of articles but also as a way to identify and count key events, such as social protests. Newspapers can be an invaluable source of public information even if they do not cover all events (i.e., selection bias) or do not report all information on the events covered (i.e., description bias). In addition, these types of bias may vary by geographic area or historical period. Although major newspapers have subject indexes, these are not always organized to be useful for social research purposes. Especially in countries with a free press, newspapers can be a way to measure social events across time. In particular, “for many historical and comparative research designs, newspapers remain the only source of data on protest events” (Earl et al., 2004:76).

### EXAMPLE BOX 5

#### Existing Statistics, Androgynous First Names, and Collective Behavior

An androgynous first name is one that can be for either a girl or boy without clearly marking the child’s gender. Some argue that the feminist movement decreased gender marking in a child’s name as part of its broader societal influence to reduce gender distinctions and inequality. Others observe that gender remains the single most predominant feature of naming in most societies. Even when racial groups or social classes invent distinctive new first names, the gender distinctions are retained.

Lieberson et al. (2000) examined existing statistical data in the form of computerized records from the birth certificates of 11 million births of White children in the state of Illinois from 1916 to 1989. They found that androgynous first names are rare (about 3 percent) and that there has been a very slight historical trend toward androgyny, but only in very recent years. In addition, parents give androgynous names to girls more than to boys, and gender segregation in naming is unstable (i.e., a name tends to lose its androgynous meaning over time). The authors noted that the way parents name children mimics a pattern of collective behavior found to operate in another research area: the racial segregation of neighborhoods. Change in residence is unequal among races with less movement by the dominant group; the less powerful group moves to occupy areas that the dominant group has abandoned; and integration is unstable with new segregation reappearing after some time.

single most valuable source of statistical information about the United States is *The Statistical Abstract of the United States*. It has been published annually (with a few exceptions) since 1878 and is available in all libraries and on the Internet. It is a selected compilation of the many reports and statistical tables of data that U.S. government agencies collect. It has summary information from hundreds of more detailed government reports, which could be examined further. With 1,400 charts, tables, and statistical lists from hundreds of government and private agencies, it is difficult to grasp all it contains



## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

until you spend time skimming through its many tables.

Most governments publish similar statistical information such as *Yearbook Australia*, *Canada Yearbook*, *New Zealand Official Yearbook*, and in the United Kingdom, the *Annual Abstract of Statistics*.<sup>13</sup> Many other nations also publish books with historical statistics.

Locating government statistical documents is an art in itself and some publications exist solely to assist the researcher: the *American Statistics Index: A Comprehensive Guide and Index to the Statistical Publications of the U.S. Government* and *Statistics Sources: A Subject Guide to Data on Industrial, Business, Social Education, Financial and Other Topics for the United States and Internationally*.<sup>14</sup> The United Nations and international agencies such as the World Bank have their own publications with statistical information (e.g., literacy rates, percentage of the labor force working in agriculture, birth rates) for various countries, for example, the *Demographic Yearbook*, *UNESCO Statistical Yearbook*, and *United Nations Statistical Yearbook*.

Other publications offer sources of data on specialized topics. For example, there are publications that contain social background, career, and other biographical information on famous individuals identified as important by some criteria. These publications depend on voluntary information provided by those deemed important. Another source of information covers businesses or their executives.<sup>15</sup> Finally, there are publications that specialize in information about politics, voting, and politicians (see Expansion Box 7, *Specialized Publications That Provide Social Data*, for source publications covering the United States).

### SECONDARY ANALYSIS OF SURVEY DATA

Secondary analysis is a special case of existing statistics. It statistically analyzes survey data originally gathered by someone else as opposed to primary research (e.g., experiments, surveys, and content analysis) that collects quantitative data. During the past two decades, many more social scientists have conducted secondary analysis as more data have become available. It is relatively inexpensive; it

permits comparisons across groups, nations, or time; it facilitates replication; and it permits asking about issues not considered by the original researchers, such as using a health survey of teens to study religion (see Example Box 6, *Secondary Data Analysis, Answering New Questions from Old Data*).

Large-scale survey data collection can be very expensive and difficult to conduct. For most researchers, the cost and time required for a major national survey that uses rigorous techniques are prohibitive. Fortunately, the organization, preservation, and dissemination of major survey data sets have improved. Today, archives of past surveys are open to researchers.

The Inter-University Consortium for Political and Social Research (ICPSR) at the University of Michigan is the world's major archive of social science data. More than 17,000 survey research and related sets of information are stored and made available to researchers at modest costs. Various centers in the United States and other nations also hold survey data.<sup>16</sup>

A widely used source of survey data for the United States is the General Social Survey (GSS). The National Opinion Research Center at the University of Chicago conducted the survey about every other year since 1973. In recent years, it has covered other nations as well. The data are made publicly available for secondary analysis at a low cost<sup>17</sup> (see Expansion Box 8, *The General Social Survey*).

### Limitations

Despite the growth and popularity of secondary data analysis and existing statistics research, they have limitations. The use of such techniques is not trouble free simply because a government agency or research organization gathered the data.

One danger is that the secondary data or existing statistics may be inappropriate for your research question. Before proceeding, you need to consider the units in the data (e.g., types of people, organizations), the time and place of data collection, the sampling methods used, and the specific issues or topics covered in the data such as the census. For example, you want to examine racial-ethnic

**EXPANSION BOX 7****Specialized Publications That Provide Social Data****PUBLISHED INFORMATION SOURCES ON FAMOUS INDIVIDUALS**

*Who's Who in America* is a popular biographic source that has been published since 1908. It lists the name, birth date, occupation, honors, publications, memberships, education, positions held, spouse, and children's names for those included. Specialized editions are devoted to regions of the United States (e.g., *Who's Who in the East*), to specific occupations (e.g., *Who's Who in Finance and Industry*), and to specific subgroups (e.g., women, Jews, African Americans).

*Dictionary of American Biography* is a more detailed listing on fewer people than *Who's Who*. It began in 1928 and has supplements to update information. For example, Supplement 7 lists 572 people and devotes about a page to each. It has details about careers, travels, the titles of publications, and relations with other famous people.

*Biographical Dictionaries Master Index* is an index listing names in the various *Who's Who* publications and many other biographic sources (e.g., *Who's Who in Hockey*). If a researcher knows a name, the index tells where biographic information can be found for the person.

**SOURCES ON BUSINESSES AND COMPANIES**

*Dun and Bradstreet Principal Industrial Businesses* is a guide to approximately 51,000 businesses in

135 countries with information on sales, number of employees, officers, and products.

*Who Owns Whom* comes in volumes for nations or regions (e.g., North America, the United Kingdom, Ireland, and Australia). It lists parent companies, subsidiaries, and associated companies.

*Standard and Poor's Register of Corporations, Directors and Executives* lists about 37,000 U.S. and Canadian companies. It has information on corporations, products, officers, industries, and sales figures.

**SOURCES ON POLITICAL ISSUES (UNITED STATES)**

*Almanac of American Politics* is a biannual publication that includes photographs and a short biography of U.S. government officials. Committee appointments, voting records, and similar information are provided for members of Congress and leaders in the executive branch.

*America Votes: A Handbook of Contemporary American Election Statistics* contains detailed voting information by county for most statewide and national offices. Primary election results are included down to the county level.

*Vital Statistics on American Politics* provides dozens of tables on political behavior, such as the campaign spending of every candidate for Congress, their primary and final votes, ideological ratings by various political organizations, and a summary of voter registration regulations by state.

tensions between Latinos and Anglos across the United States but have only secondary data that includes the Pacific Northwest and New England states. In this situation, you should reconsider the question or the use of data.

A second danger is that you must understand the substantive topic to use the data. Because they are easily accessible, you might have data but know very little about a topic. As a result, you make erroneous assumptions or false interpretations about the results. Before using any data, you should study details of the substantive topic. For example, if you use data on high school graduation rates in Germany but you do not know much about the German

secondary education system with its distinct academic and vocational tracks and assume the German and U.S. system are the same, you can easily make serious errors in interpreting results.

A third danger is to quote statistics in excessive detail to give others an impression of scientific rigor. This can lead to the **fallacy of misplaced concreteness**, which occurs when someone gives a false impression of precision by quoting statistics in more

**Fallacy of misplaced concreteness** Use of too many digits in a quantitative measure in an attempt to create the (mis)impression that data are accurate.

**EXAMPLE BOX 6****Secondary Data Analysis, Answering  
New Questions from Old Data**

To perform secondary data analysis, researchers can use already collected survey data to address new research questions unrelated to the survey's original purpose. Uecker et al. (2007) used data from a health survey to study religion. Data were from National Longitudinal Study of Adolescent Health, a school-based three-part panel survey on health and related social behaviors. The authors used data from the first panel of the survey, a random sample in 1994–1995 of 132 schools with 20,745 U.S. adolescents in grades 7–12 and from the third part of the panel, interviews in 2001–2002 with 15,197 of the original respondents (who were then aged 18–25). The authors' interest was to explain declines in religious involvement that occur as young adults move from adolescence to adulthood, not in health. However, the health survey had questions about religious involvement, importance of religion in one's life, and feelings about organized religion. The authors used these three questions as their dependent variable measure. Past research had explained the decline in religion during the young adult years as being due to the secularizing effects of going to college. After statistically analyzing the data, the authors found that people who went to college remained as religious as those who did not go. The authors found that cohabitation, non-marital sex, and drug and alcohol use among the young people reduced the importance of religion in the young person's life, not whether they attended college, and that contrary to what people had thought, higher education itself had little effect on religious belief.

detail than warranted and “overloading” the details. For example, existing statistics report that the population of Australia is 19,169,083, but it is better to say that it is a little more than 19 million. You might calculate the percentage of divorced people as 15.65495 in a secondary data analysis of the 2000 General Social Survey, but it is better to report that about 15.7 percent of people are divorced.<sup>18</sup>

**Units of Analysis and Variable Attributes.** A common problem in existing statistics is finding the

**EXPANSION BOX 8****The General Social Survey**

The General Social Survey (GSS) is the best-known set of survey data used by social researchers for secondary analysis. The mission of the GSS is “to make timely, high quality, scientifically relevant data available to the social science research community” (Davis and Smith, 1992:1). It is available in many computer-readable formats and is widely accessible for a low cost. Neither datasets nor codebooks are copyrighted. Users may copy or disseminate them without obtaining permission. You can find results using the GSS in more than 2,000 research articles and books.

The National Opinion Research Center (NORC) has conducted the GSS almost every year since 1972. A typical year's survey contains a random sample of about 1,500 adult U.S. residents. A team of researchers selects some questions for inclusion, and individual researchers can recommend questions. The Center repeats some questions and topics each year, includes some on a four- to six-year cycle, and adds other topics in specific years. For example, in 1988, the special topic was religion, and in 1990, it was intergroup relations.

Interviewers collect the data through face-to-face interviews. The NORC staff carefully selects interviewers and trains them in social science methodology and survey interviewing. About 120 to 140 interviewers work on the GSS each year. About 95 percent are women, and most are middle-aged. The NORC recruits bilingual and minority interviewers. Interviewers are race-matched with respondents. Interviews are typically 90 minutes long and contain approximately 500 questions. The response rate has been 71 to 79 percent. The major reason for nonresponse is a refusal to participate.

The International Social Survey Program conducts similar surveys in other nations. Beginning with the German ALLBUS and British Social Attitudes Survey, participation has grown to include Australia, Austria, Italy, Hungary, Ireland, Israel, the Netherlands, Switzerland, Poland and others. The goal is to conduct on a regular basis large-scale national general surveys in which some common questions are asked across cooperating nations.

appropriate units of analysis. Many statistics are published for aggregates, not the individual. For example, a table in a government document has

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

information (e.g., unemployment rate, crime rate) for a state, but the unit of analysis for the research question is the individual (e.g., “Are unemployed people more likely to commit property crimes?”). The potential for committing the ecological fallacy is very real in this situation. It is less of a problem for secondary survey analysis because we can obtain raw information on each respondent from archives.

A related problem involves the categories of variable attributes used in existing documents or survey questions. This is not a problem if organizations that gathered the initial data used many highly refined categories. The problem arises when the organizations collected the original data in broad categories or ones that do not match the needs of current research. For example, you are interested in people of Asian heritage. If the racial and ethnic heritage categories in a document are White, Black, and Other, you have a problem. The Other category includes people of Asian and other heritages. Sometimes organizations gather information in refined categories but publish it only in broad categories. You need to dig more deeply to discover whether the organization collected refined information.

**Validity.** Validity problems can occur when your theoretical definition does not match that of the government agency or organization that collected the information. Official policies and procedures specify definitions for official statistics. For example, you define a work injury as including minor cuts, bruises, and sprains that occur on the job, but the official definition in government reports includes only injuries that require a visit to a physician or hospital. Many work injuries that you define as relevant will not be included in official statistics. Another example occurs when you define as unemployed people who would work if a good job were available, who have to work part-time when they want full-time work, and who have given up looking for work, but the official definition of unemployed includes only those who are actively seeking work (full- or part-time). The official statistics exclude those whom you define as unemployed. In both cases, your definition differs from that in official statistics.

Another validity problem arises when you rely on official statistics as a proxy for a construct. This is necessary because you cannot collect original

data. For example, you want to know how many people are victims of hate crimes, so you use police statistics on hate crime as a proxy, but the measure is not entirely valid. Many victims do not report hate crimes to the police, and official reports do not always reveal all that occurred (see Expansion Box 9, Official Statistics on Hate Crime, Slow Improvements in Accuracy).

Perhaps you want to measure marriages “forced” by a premarital pregnancy. You can use the date of marriage and the date of the birth of a child in official records to estimate whether such a marriage occurred. This does not tell you that pregnancy was the motivation for the marriage, however. A couple may have planned to marry and the pregnancy was irrelevant, or the pregnancy may have been unknown at the date of marriage. Likewise, some marriages that show no record of a birth could have been forced by a false belief in pregnancy, or a pregnancy that ended in a miscarriage or abortion instead of a birth. In addition, a child might be conceived after the date of marriage, but be born very prematurely. If you measure forced marriages as those in which a child was born less than nine months after a marriage date, some will be mislabeled, thereby lowering your study’s validity.

A third validity problem arises because you lack control over how information is collected. Ordinary people who work in bureaucracies collect information that appears in official government reports. You depend on these people to collect, organize, report, and publish data accurately. Systematic errors in collecting the initial information (e.g., census workers who avoid poor neighborhoods and make up information or people who put a false age on a driver’s license), in organizing and reporting information (e.g., a police department that is sloppy about filing crime reports and loses some), and in publishing information (e.g., a typographical error in a table) all reduce measurement validity.

Such a problem happened in U.S. statistics regarding the number of people permanently laid off from their jobs. A university researcher reexamined the methods used to gather data by the U.S. Bureau of Labor Statistics and found an error. Data on permanent job losses came from a survey of 50,000 people, but the government agency failed to adjust for a high survey nonresponse rate. The

**EXPANSION BOX 9****Official Statistics on Hate Crime, Slow Improvements in Accuracy**

Government statistics on crime is one of the many types of existing statistics frequently used in social research. The Uniform Crime Reporting (UCR) program, operating since 1920, is the most-used source of national crime statistics in the United States. Each state and local law enforcement agency sends its crime statistics (i.e., crimes reported to police and arrests made) for most major crimes to the Federal Bureau of Investigation. In 1973, the Department of Justice added a second source of crime data, the National Crime Victimization Survey (NCVS). It is an annual survey conducted with a representative sample of 49,000 households. The survey asks household members whether anyone over the age of 12 had been a crime victim. In the 1980s, a new program supplemented the simple crime counts from the UCR. The National Incident-Based Reporting System (NIBRS) includes many more details about crime circumstances (e.g., location, participants, time). It now covers about 15 percent of the U.S. population, but is slowly expanding.

In 1990, the United States enacted a new hate crime law in which crimes committed that include bias or prejudice regarding a victim's race, religion, ethnicity, or sexual orientation result in added penalties. The FBI publishes a summary of such crimes in its annual report *Hate Crime Statistics*, the primary source of national information on such crime. The accuracy of reporting this new crime illustrates some complexities with official statistics.

After enactment of the national hate crime law, the FBI trained local law officials on enforcing the new

law. In 1991, only 29 percent of the U.S. population had a law enforcement agency that had participated in this training program. By 1999, this had risen to 85 percent of the population. However, a majority of participating agencies report zero hate crimes each year. Since 1994, about 85 percent of trained law enforcement agencies reported zero hate crimes. These zero reports may not be entirely accurate. One study of locations with zero reports discovered that 37 percent of them had hate crimes, but a breakdown occurred in the reporting system. Large regional reporting differences exist. Southern states, with the highest general crime rates, have the lowest reports of hate crimes. Some studies suggest that local attitudes about hate crimes influence the willingness of victims to report them to the police. In addition, not all local police agencies take the hate crime violations equally seriously. Starting in 2000, the National Victimization Survey began to add hate crime questions.

In summary, a dozen years after hate crime legislation was enacted, such crime data are limited. Reporting accuracy is uneven by geographic area and police agency. The data collection, which provides a basic understanding about the numbers and types of hate crimes, is improving slowly. Researchers who want to study hate crime seriously need to combine official national reports with other sources of information that contain more detail and focus on local geographic areas (see Nolan, Akiyama, and Berhanu, 2002).

corrected figures showed that instead of a 7 percent decline in the number of people laid off between 1993 and 1996, as had been first reported, there had been no change.<sup>19</sup>

**Reliability.** Reliability problems can plague existing statistics research; they occur when official definitions or the method of collecting information changes over time. Official definitions of work injury, disability, unemployment, and the like change periodically. Even if we learn of such changes, consistent measurement over time is impossible. For example, during the early 1980s, the method for

calculating the U.S. unemployment rate changed. Previously, the government had calculated the unemployment rate as the number of unemployed persons divided by the number in the civilian workforce. The new method divided the number of unemployed by the civilian workforce plus the number of people in the military. Likewise, when police departments computerize their records, the number of crimes reported appears to increase, simply because of improved record-keeping.

Equivalence reliability can also be a problem. For example, a measure of crime across a nation depends on each police department's providing

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

accurate information. If departments in one region of a country have sloppy bookkeeping, the measure reported loses equivalence reliability. Likewise, studies of police departments suggest that political pressures to increase arrests are closely related to the number of arrests. For example, political pressure in one city may increase arrests (e.g., a crackdown on crime) whereas pressures in another city may decrease arrests (e.g., to show a drop in crime shortly before an election in order to make officials look better).

Representative reliability can be a problem in official government statistics. For example, the U.S. Bureau of Labor Statistics found a 0.6 percent increase in the female unemployment rate after it used gender-neutral measurement procedures. Until the mid-1990s, interviewers asked women only whether they had been keeping house or doing something else. Researchers categorized women who answered “keeping house” as being housewives, not as being unemployed, even if the women had been seeking work. Once they asked women the same question as men, “Were you working or doing something else?” many women reported not working but doing “something else” such as looking for work. This shows the importance of methodological details in how officials create government statistics.

Official statistics allow for international comparisons but national governments collect data differently and the quality of data collection varies. For example, in 1994, the official unemployment rate reported for the United States was 7 percent, 2.9 percent in Japan, and 12 percent in France. If the nations defined and gathered data the same way, including rates of discouraged workers and involuntary part-time workers, the rates would have been 9.3 percent for the United States, 9.6 percent for Japan, and 13.7 percent for France. To evaluate the quality of official government statistics, *The Economist* magazine asked a team of 20 leading statisticians to evaluate the statistics of thirteen nations based on freedom from political interference, reliability, statistical methodology, and coverage of topics. The top five nations in order were Canada, Australia, Holland, France, and Sweden. The United States tied for sixth with Britain and Germany. The quality of U.S. statistics suffered from being highly decentralized, having fewer

statisticians than any other nation, and experiencing politically motivated cutbacks on the range of data collected.

Data collected internationally can be controversial. The International Labor Organization of the United Nations reported that the official statistics of total economic activity for several nations are inaccurate because they exclude the sex industry. In some countries (especially Thailand and the Philippines), millions of workers (primarily young women) are employed and billions of dollars in revenue are generated from prostitution and the sex industry. This has a large impact on the economy, but it does not appear in any official reports or statistics.<sup>20</sup>

**Missing Data.** One problem that plagues researchers who use existing statistics and documents is that of missing data. Sometimes the data were collected but lost. More frequently, the data were never collected. The data may be missing because researchers and officials in government agencies decided not to collect information. Those who decide what to collect may not collect what later researchers will need in order to address new questions. Government agencies start or stop collecting information for political, budgetary, or other reasons. For example, during the early 1980s, cost-cutting measures by the U.S. federal government stopped the collection of information that social researchers found valuable. Missing information is a problem especially when researchers cover long periods. For instance, someone interested in studying the number of work stoppages and strikes in the United States can obtain data from the 1890s to the present except for a 5-year period after 1911 when the federal government did not collect the data.

## ISSUES OF INFERENCE AND THEORY TESTING

You need to take extra care when inferring causality or testing a theory based on nonreactive data. It is difficult to establish temporal order and eliminate alternative explanations with nonreactive and unobtrusive measures. In content analysis, you cannot generalize from the content to its effects on those

## NONREACTIVE RESEARCH AND SECONDARY ANALYSIS

who read the text, but can use the correlation logic of survey research only to show an association among variables.

### **ETHICAL CONCERNS**

Ethical concerns are not at the forefront of most nonreactive research because the people you study are not directly involved. The primary ethical concern is the privacy and confidentiality of using information that someone else gathers. Another larger ethical issue is that official statistics are social and political products. Some researchers or official agencies gather data based on implicit theories and value assumptions. Official measures or statistics can be the objects of political conflict and a way to push policy in certain political directions. Once government agencies define a measure as official, it can influence public policy and lead to outcomes that would be different had an alternative but equally valid measure been used. For example, political activism during the Great Depression of the 1930 simulated the collection of information on many social conditions (e.g., the number of patients who died while in public mental hospitals). Before the political activism of the time, governments and others did not see the conditions as sufficiently important to warrant public attention. Likewise, information on the percentage of non-White students enrolled in U.S. schools at various ages is available only since 1953 and for various non-White races only since the 1970s. Earlier, such information was not salient for public policy.

The collection of official statistics can stimulate public attention toward an issue, and public concern about a social issue can stimulate the collection of new official statistics. For example, drunk driving became a public issue only after government agencies started to maintain statistics on the number of automobile accidents in which alcohol was a factor.

Political and social values influence decisions about which statistics government agencies collect. The design and collection of most official statistics is for top-down administrative planning purposes. The data may not conform to your purposes or the purposes of people who disagree with the thinking

of bureaucratic decision makers. For example, a government agency measures the number of tons of steel produced, miles of highway paved, and the average number of people in a household. Information on other conditions such as drinking-water quality, time needed to commute to work, stress related to a job, and number of children needing child care may not be collected because political officials consider it to be unimportant. In many countries, officials see gross national product (GNP) as a critical measure of societal progress, but GNP ignores noneconomic aspects of social life (e.g., time spent playing with one's children) and types of work (e.g., housework) that are free. The information available reflects the outcome of political debate and the values of officials who decide which statistics to collect.<sup>21</sup>

### **CONCLUSION**

In this chapter, you read about several types of nonreactive research techniques. They are ways to measure or observe aspects of social life without affecting those who are being studied. They result in objective, numerical information that you can analyze to address research questions. You can use the techniques in conjunction with other types of quantitative or qualitative social research to address a large number of questions.

As with any form of quantitative data, we need to be concerned with measurement issues. It is easy to take available information from a survey or government document, but this does not mean that it measures the construct of interest to us.

You should be aware of two potential problems in nonreactive research. First, the availability of existing information restricts the questions that we can address. Second, the nonreactive variables often have weak validity because they do not measure the construct of interest. Although existing statistics and secondary data analysis are low-cost research techniques, the researcher lacks control over, and substantial knowledge of, the data collection process. This potential source of errors means that researchers need to be especially vigilant and cautious.